

Enhancing EPHT with Satellite-Driven PM_{2.5} Exposure Modeling and Epidemiology – Year 4 (NCE) Report



Yang Liu, Ph.D.
Rollins School of Public Health
Emory University



September 25, 2013
St. Paul, MN



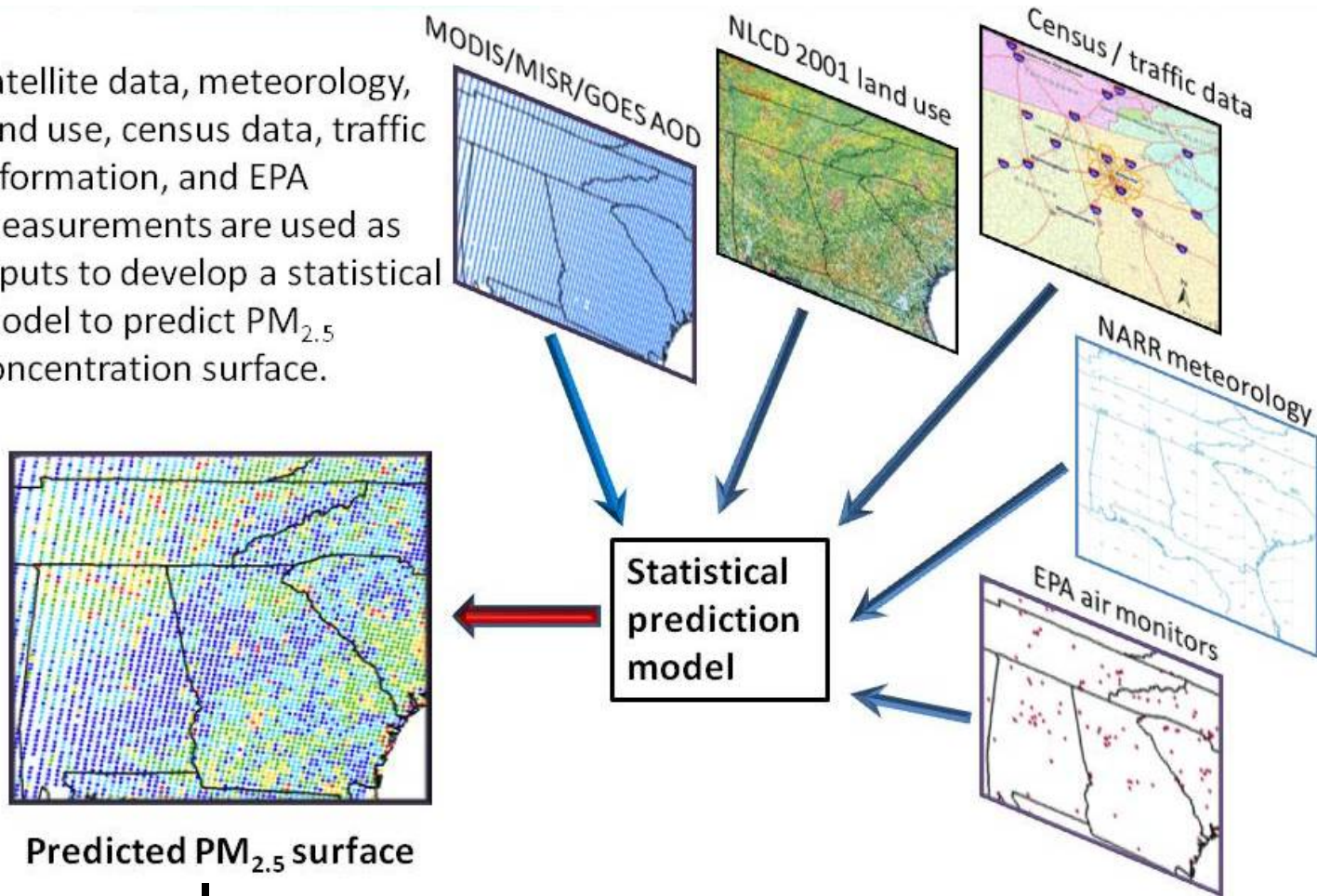
- ❑ Emory: Yang Liu (PI), Jeremy Sarnat, Mitch Klein, and Xuefei Hu
- ❑ MSFC/USRA: Dale Quattrochi, Bill Crosson, Mohammad Al-Hamdan, Maury Estes, Sue Estes, Sarah Hemmings, and Gina Wade
- ❑ CDC: Judy Qualters, Paul Garbe, Helen Flowers, and Ambarish Vaidyanathan, Heather Strosnider, and Erika Rees

*Sponsored by NASA Applied Science Program grant
NNX09AT52G under J. Haynes and Sue Estes*

- ❑ Extend the spatial coverage of the $PM_{2.5}$ indicators in CDC Tracking Network with NASA Earth observations
- ❑ Explore the utility of AOD based county level $PM_{2.5}$ health indicators for public health surveillance
- ❑ Evaluate satellite $PM_{2.5}$ estimates as a alternative exposure data source in environmental epidemiologic studies

- ❑ A: integration of Earth science data (Year 1)
 - ❑ Spatially and temporally match various data sources to a defined master grid in study domain
- ❑ B: PM_{2.5} exposure modeling (Year 2)
 - ❑ Develop spatial statistical models to estimate PM_{2.5} concentrations and compare with existing Tracking datasets and independent field measurements
- ❑ C: PM_{2.5} health effects modeling (Year 3 and NCE)
 - ❑ Tracking evaluation and website launch
 - ❑ Associate model estimated PM_{2.5} concentrations with cardiorespiratory ED visits in an epidemiological model, and compare effects with conventional methods

Satellite data, meteorology, land use, census data, traffic information, and EPA measurements are used as inputs to develop a statistical model to predict $PM_{2.5}$ concentration surface.



Prospective
sampling

Tracking portal
dissemination

Epidemiological
modeling

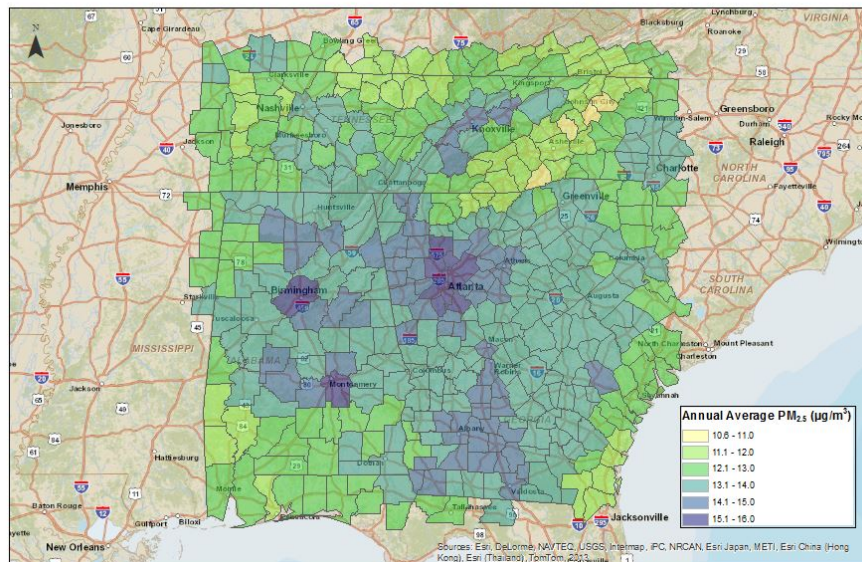
Milestones	Deadline	Team	Status
Dissemination Through the Tracking Network	09/2013	CDC	Complete
Training and outreach	09/2013	Emory/CDC	Complete
Academic conferences	09/2013	Emory/CDC	ATS, URISA, ISEE
Epidemiological modeling	09/2013	Emory	Complete
Manuscripts	09/2013	Emory /CDC	ongoing
Final report	12/2013	Emory	In preparation

Final ARL: 8 (proposed ARL: 7)

‘Clear communication’
information about PM_{2.5},
remote sensing, monitored
and modeled sources of
PM_{2.5} data

Remotely Sensed Annual Average PM_{2.5} Estimates, 2007

Remote Sensing data are available
for the Southeastern United States only.



National Environmental Public Health Tracking

- Air Quality
- Tracking Air Quality
- Monitor + Modeled Air Data
- Health Impacts of Fine Particles in Air
- Related Links
- Air Quality Indicators
- Outdoor Air Communication Tools
- Search Air Quality Data

Tracking Links | **Environments** | **Health Effects** | **Population Health** | **Info by Location**

Quick Links

- Air and Health
- Air Monitoring in the US
- Air Contaminants
- Air Toxics **NEW**
- Remote Sensing data **NEW**

Tracking Success Stories

- Colorado
- Florida
- Maine
- Maryland
- Massachusetts
- New York
- Washington

Atmospheric Remote Sensing: Modeled PM_{2.5}

Atmospheric remote sensing can be used to measure levels of some air pollutants. Remote sensing data come from satellites. These data can be used in combination with other data to help us better understand when and where air pollution is happening. This is important because air pollution can cause health problems. Knowing more about when and where air pollution is happening can help public health officials and others do more to protect our health. Read more about the health problems related to air pollution [here](#).

Compliance with air pollution standards usually is measured through the use of EPA's [Air Quality System](#) monitors. These monitors are on the ground and are placed around the country, mainly in large cities. Data from these monitoring stations are considered the "gold standard" for determining outdoor air pollution. However, this information is limited because the monitoring stations are usually near big cities and may take air samples only every three days or during periods when air pollution is very high. Read more about air monitoring [here](#). Using remote sensing data from satellites can help fill in the gaps that exist from air monitors on the ground.

Although atmospheric remote sensing data can help estimate air pollution levels, these data have limitations especially if used on their own. Satellite data are not always available. For example, it is nearly impossible to collect satellite data on a cloudy day. Clouds can interfere with the satellite's ability to collect data which can cause a gap in the information that comes from them. This is one reason why atmospheric remote sensing data should be used in addition to monitoring and modeled air data.

The National Aeronautics and Space Administration (NASA) provides atmospheric sensing data from their satellites for this project. Scientists from CDC, NASA, and Emory University are working together to determine how these data can be used with other air pollution monitoring data to measure fine particulate matter in outdoor air. Fine particulate matter is also called PM_{2.5}. Read more about PM_{2.5} [here](#).

The Tracking Network is now providing estimates of annual average PM_{2.5} concentrations using remote sensing data. Currently, data are available only for Alabama, Georgia, and parts of South Carolina, Tennessee, North Carolina, Florida, and Virginia.

Home

- About Tracking Program
- State & Local Tracking Portals
- Success Stories
- Indicators & Data
- Secure Portal
- Print page
- Bookmark and share
- CDC on Facebook
- CDC on Twitter

Tracking Hot Topics

- Download the Extreme Heat: Prevent Heat-Related Illness Widget
- Tips for Preventing Heat-Related Illness
- Stay Healthy and Safe in Hot Weather PSA
- Recognizing, Preventing, and Treating Heat-Related Illness Online Training
- View our Tracking Success Stories to learn how Tracking is making a difference across the U.S.

Resources

- Communication Tools
- Training
- Join our List-serv
- Document Library
- Quick Reports
- Technical Notes

Contact Us:

- Centers for Disease Control and Prevention
- 1600 Clifton Rd
- Atlanta, GA 30333
- 800-CDC-INFO (800-232-4636)
- TTY: (888) 232-6348
- New Hours of Operation
- 8am-8pm ET/Monday-Friday

Can download 2001-2007 AOD-based PM_{2.5} data for the southeast U.S. at county-level



National Environmental Public Health Tracking Program Academic Partners of Excellence

A Summer Webinar Series Featuring Research by Tracking Program Partner Institutions

Part VII – Assessment of Remotely Sensed PM_{2.5} Estimates for the Tracking Network

Wednesday, August 28, 2013

1:30 – 2:30 p.m. EST

Chamblee Campus, Building 107

Conference Room 3A

Join the webinar: <https://www1.gotomeeting.com/register/391731697>

Register for the webinar to get the call-in information.

Presenters

Yang Liu, Ph.D.

Emory University

Rollins School of Public Health

Rish Vaidyanathan

Centers for Disease Control and Prevention

National Environmental Public Health Tracking Program

1. The application of satellite remote sensing data in a time-series study of asthma exacerbation in Metro Atlanta. **Conference of ISEE, ISES and ISIAQ**, Basel, Switzerland, August 19–23, 2013.
2. A time series analysis of PM_{2.5} concentrations in the southeastern US using MAIAC AOD in a two-stage spatial statistical model. **Conference of ISEE, ISES and ISIAQ**, Basel, Switzerland, August 19–23, 2013.
3. Enhancing EPHT with satellite-driven PM_{2.5} exposure modeling and epidemiology. **URISA's Fourth GIS in Public Health Conference**, Miami, FL, June 17–20, 2013.
4. Estimating ground-level PM_{2.5} concentrations in the southeastern United States using MAIAC AOD retrievals and a two-stage model. **American Thoracic Society International Conference**, Philadelphia, PA, May 17–22, 2013.
5. Estimating ground-level PM_{2.5} concentrations in the southeastern US using MAIAC AOD retrievals, **ISES Annual Meeting**, Seattle, WA, October 30, 2012.

1. Hu X*, Waller LA, Al-Hamdan MZ, Crosson WL, Estes Jr MG, Estes SM, Quattrochi DA, Sarnat JA, **Liu Y**. 2013. Estimating ground-level PM_{2.5} concentrations in the southeastern US using geographically weighted regression. *Environ Res*. 121:1-10.
2. Puttaswamy SJ, Nguyen H, Braverman A, Hu X, **Liu Y**. 2013. Statistical data fusion of multisensor AOD over the continental United States. *Geocarto International*. DOI: 10.1080/10106049.2013.827750.
3. Hu X*, Waller LA, Lyapustin A, Wang Y, Al-Hamdan MZ, Crosson WL, Estes MG, Estes SM, Quattrochi DA, Puttaswamy SJ, Liu Y. 2013a. Estimating ground-level PM_{2.5} concentrations in the southeastern United States using MAIAC AOD retrievals and a two-stage model. *Remote Sens Environ*. In press.
4. Kim M, Zhang X, Holt J, **Liu Y**. 2013. Spatio-temporal variations in the associations between hourly PM_{2.5} and aerosol optical depth (AOD) from MODIS sensors on Terra and Aqua. *Health*. In press.

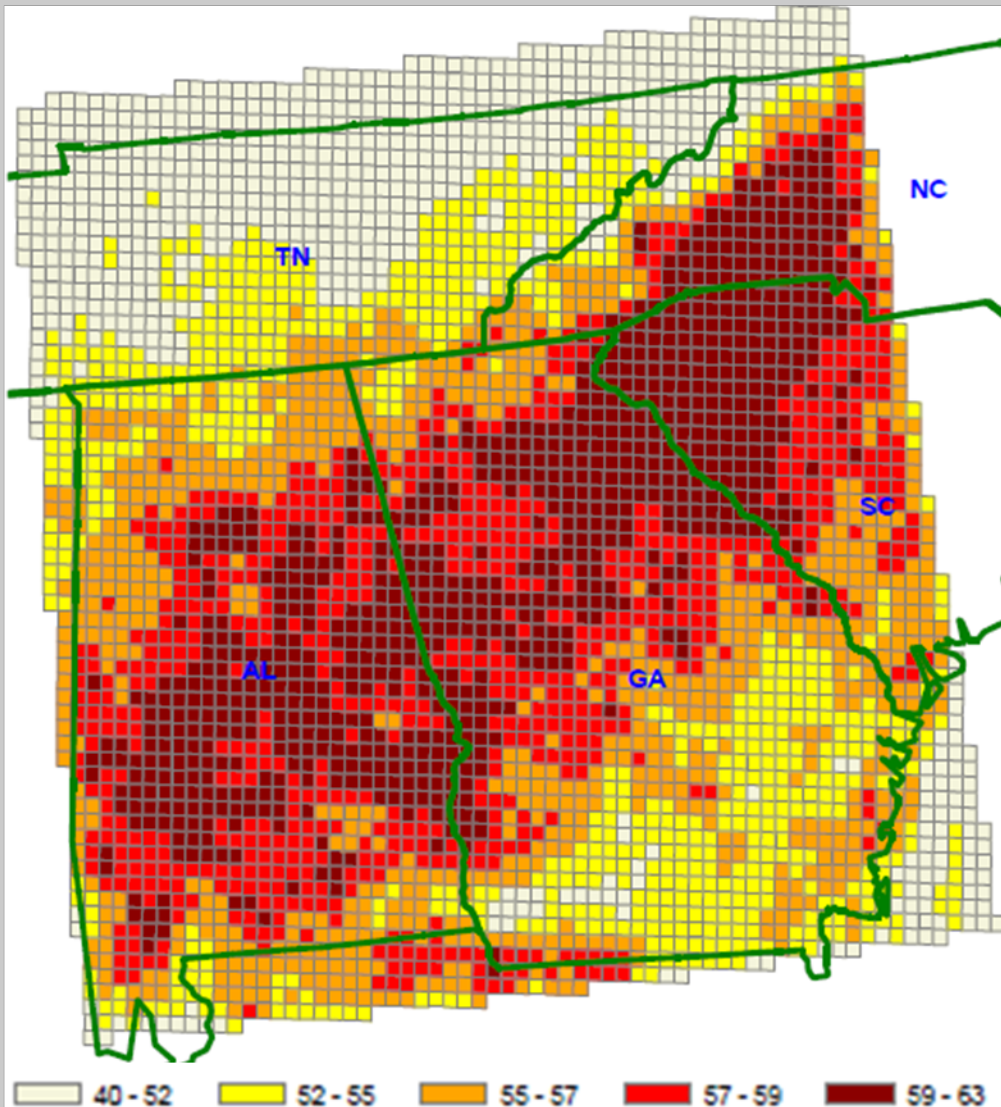
1. Chang HH, Hu X, **Liu Y**. 2013. Calibrating MODIS aerosol optical depth for predicting daily PM_{2.5} concentrations via statistical downscaling. *J Expo Anal Environ Epidemiol*. In revision.
2. Yu C*, Chen L, Zhang X, Girolamo LD, **Liu Y**. 2013. Effects of MODIS-retrieved cloud properties on PM_{2.5} levels in the southeastern United States. *Atmos Environ*. Submitted.
3. Hu X*, Waller LA, Lyapustin A, Wang Y, **Liu Y**. 2013b. 10-year spatial and temporal trends of PM_{2.5} concentrations in the southeastern US estimated using high-resolution satellite data. *Atmos Chem Phys*. Submitted.
4. One manuscript to be submitted to the Tracking special issue in *Environmental Research*.
5. One manuscript to be submitted to the ISEE/ISES/ISIAQ 2013 conference special issue in *International Journal of Environmental Research and Public Health*.
6. Maybe more.

- ❑ Federal reference methods (FRM) based monitoring data from Environmental Protection Agency's Air Quality System (AQS)
 - ❑ “Gold standard”
 - ❑ Most monitors sample every third day
- ❑ Community Multiscale Air Quality (CMAQ) model
 - ❑ A grid based numerical deterministic simulation model
- ❑ Bayesian space-time hierarchical fusion models
 - ❑ Combine air quality measurements with CMAQ predictions
 - ❑ Temporal resolution: daily
 - ❑ Version 1: hierarchical Bayesian (HB) model (available at 12 and 36 km grid resolution)
 - ❑ Version 2: Downscaler (DS) model (available at census tract centroids)

- ❑ Daily level comparison of modeled data sources of $PM_{2.5}$ against monitor data for the Southeast (2006)
 - ❑ In-sample evaluation using AQS data
 - ❑ Independent validation using measurements from the Southeastern Aerosol Research and Characterization (SEARCH) Experiment network.

- ❑ Annual level comparison to explore the feasibility of generating the county level $PM_{2.5}$ indicators (2006)
 - ❑ Annual Average
 - ❑ County level mortality benefits assessment

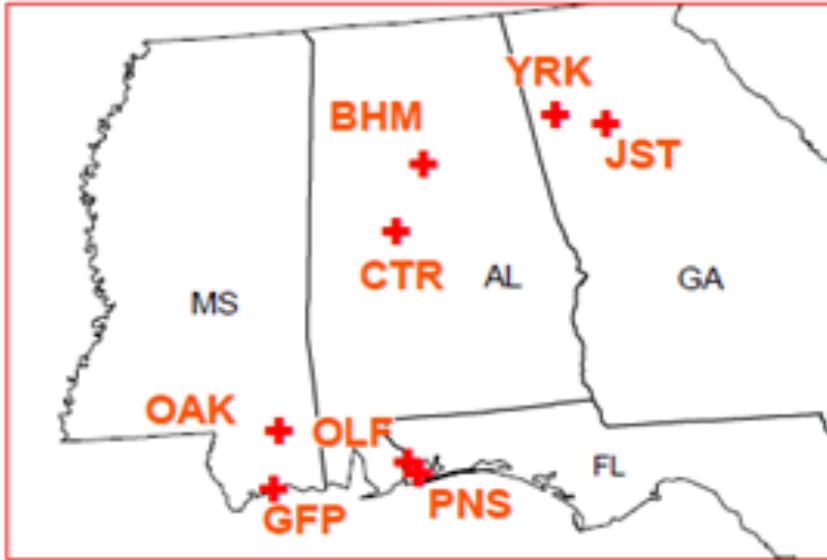
Percent of days with Remote Sensing (RS) data 2006



Data Source	Median (Range) percent of daily completeness (Annual)
AQS	32 (15-100)
Remote Sensing (AOD)	56 (40-63)
CMAQ	100 (100-100)
Downscaler	100 (100-100)
SEARCH	93 (90-95)

Completeness of AOD data varies with time of the year

- JAN – MAR has the lowest completeness
- OCT – DEC has the highest completeness



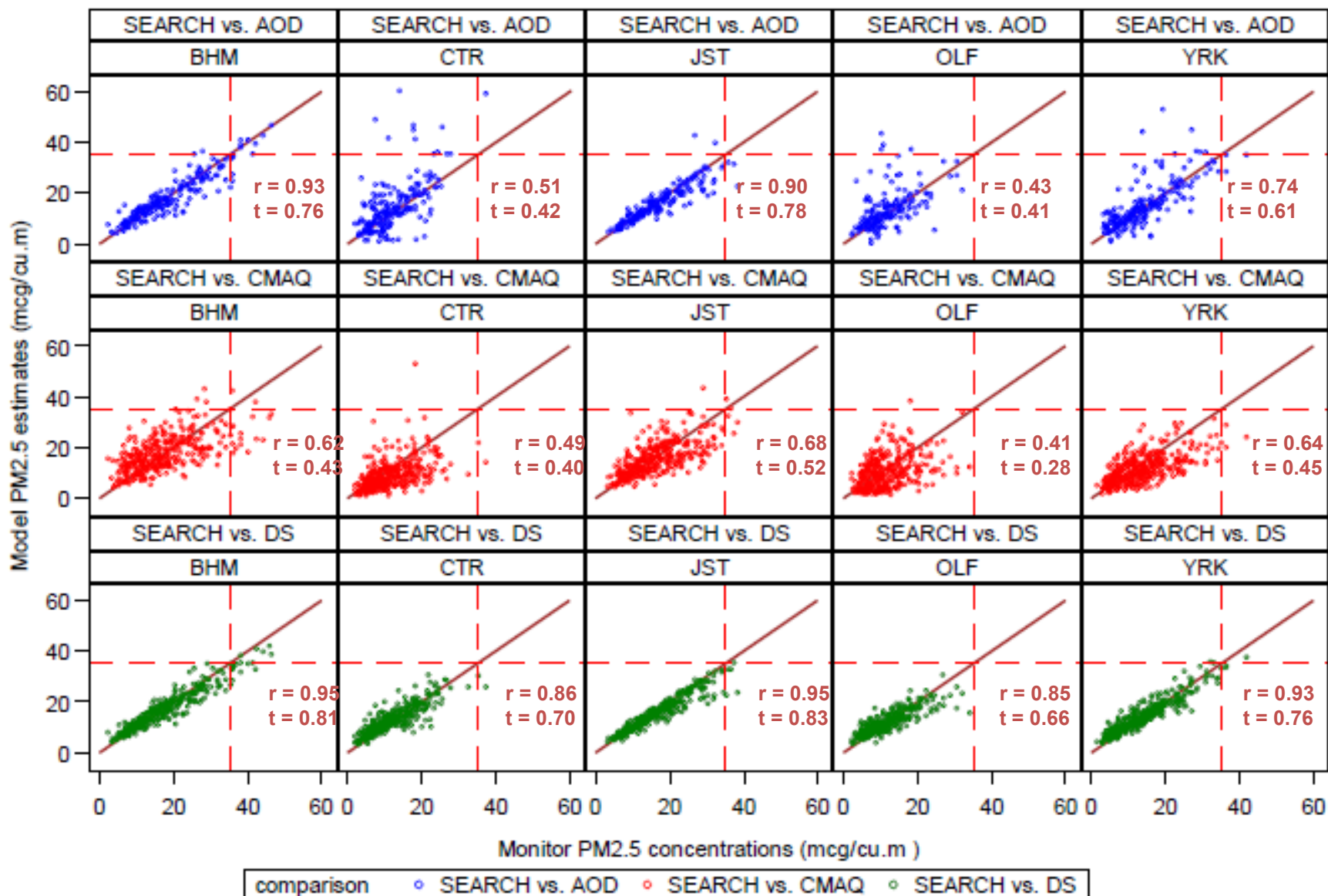
Site	Location	
CTR	Centreville, AL	X
BHM	North Birmingham, AL	X
YRK	Yorkville, GA	X
JST	Jefferson St., Atlanta, GA	
OAK	Oak Grove, MS	
GFP	Gulfport, MS	X
OLF	Outlying landing field, Pensacola, FL	
PNS	Pensacola, FL	

Southeastern Aerosol Research and Characterization (SEARCH)

Experiment network covers four states and has four urban-rural paired sites

- Developed as part of a public-private collaboration with EPRI (Electric Power Research Institute), Southern Company, and other utilities
- <http://atmospheric-research.com/>
- Primarily formed to assess air quality in the Southeast

Comparison at SEARCH Sites



Note: satellite PM_{2.5} was based on an earlier version of single stage GWR model

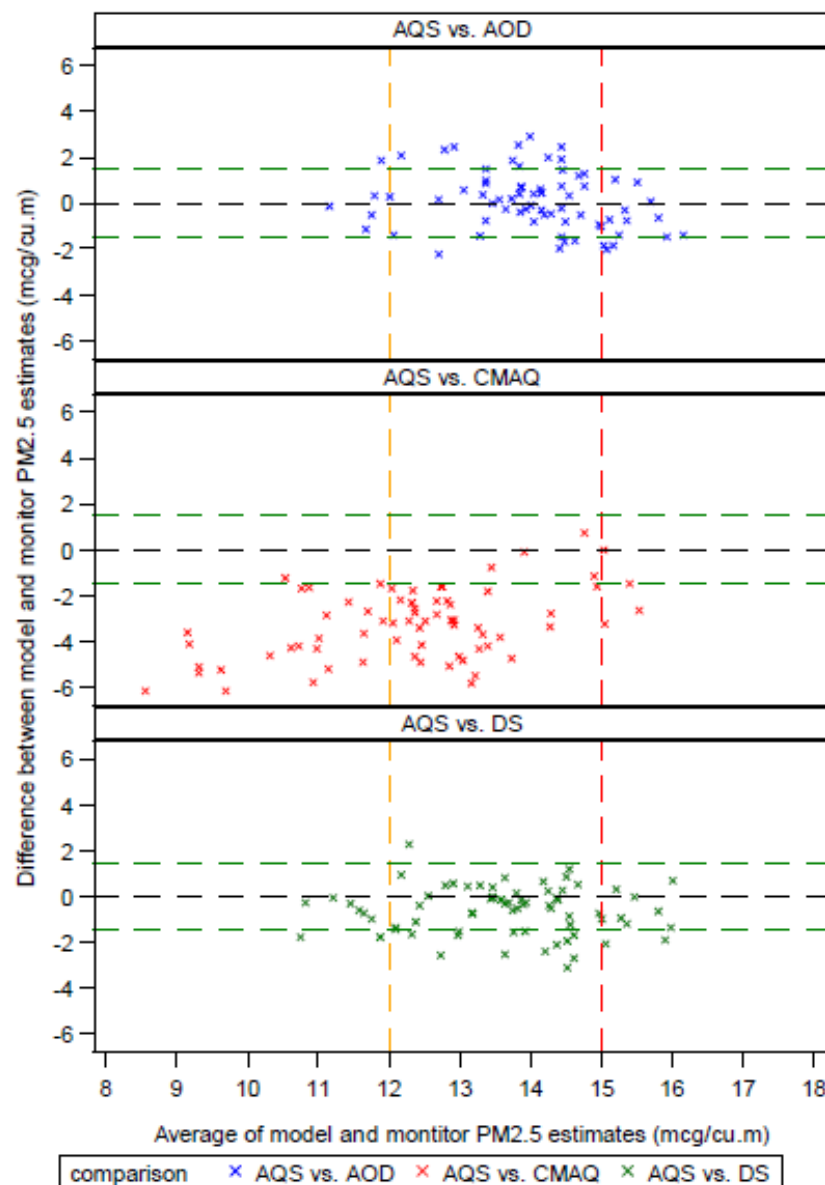
Urban sites (BHM, JST)

SEARCH Site	SEARCH vs. AOD		SEARCH vs. CMAQ		SEARCH vs. DS	
	Root Mean Squared Deviation (μg/cu.m) (RMSD)	Relative Accuracy (%) (RA)	Root Mean Squared Deviation (μg/cu.m) (RMSD)	Relative Accuracy (%) (RA)	Root Mean Squared Deviation (μg/cu.m) (RMSD)	Relative Accuracy (%) (RA)
BHM	3.6	79.4	7.1	59.3	2.8	83.6
JST	3.5	78.8	5.6	65.7	2.6	84.2

Rural sites (CTR, OLF, YRK)

SEARCH Site	SEARCH vs. AOD		SEARCH vs. CMAQ		SEARCH vs. DS	
	Root Mean Squared Deviation (μg/cu.m) (RMSD)	Relative Accuracy (%) (RA)	Root Mean Squared Deviation (μg/cu.m) (RMSD)	Relative Accuracy (%) (RA)	Root Mean Squared Deviation (μg/cu.m) (RMSD)	Relative Accuracy (%) (RA)
CTR	9.8	17.1	6.5	45.7	3.2	72.9
OLF	8.0	30.6	6.7	41.8	3.1	72.8
YRK	6.4	54.3	6.1	56.3	2.7	80.3

Bland Altman Plot: Annual Mean PM2.5



- ❑ AOD-based data compare well with AQS monitoring data
- ❑ Overall, there is good agreement between DS and AOD based predictions where AQS data are available
- ❑ Performance of AOD based $PM_{2.5}$ is relatively poor at rural locations

- ❑ PM_{2.5} data from central monitoring sites have served as the foundation for exposure assignment in epidemiologic studies
- ❑ Ground monitors are often clustered in urban centers and may not fully capture PM_{2.5} spatial contrasts over large metropolitan areas
- ❑ Satellite-based estimates of PM_{2.5} may improve exposure assignment in large population-based health studies

- ❑ Estimate associations between daily PM_{2.5} concentrations and respiratory emergency department (ED) visits in 20-county metro Atlanta for 2001-2007

- ❑ Compare results among three PM_{2.5} exposure assignment approaches
 - ❑ Centrally-located ground monitor
 - ❑ Population-weighted average of multiple monitors
 - ❑ Spatially-resolved satellite-based estimates

20-County Atlanta PM_{2.5} Results



- 101 grids over 2,556 days = 258,156 possible grid-day obs.

Metric	N	N Miss	Mean	Std Dev
Central Site	258,156	0	16.0	7.7
Central Site (matched)	141,735	116,421	17.0	8.3
Satellite	141,735	116,421	15.2	7.0

- 45% of grid-day satellite measurements missing largely due to cloud cover, therefore not at random
- PM_{2.5} estimates more similar over space than over time
 - Little difference in mean PM_{2.5} over the 20-county area
 - Range of grid cell means: 13.7 to 27.7 $\mu\text{g}/\text{m}^3$ (90th percentile = 15.7 $\mu\text{g}/\text{m}^3$)
- PM_{2.5} estimates among grid cells highly correlated
 - Temporal correlations among satellite PM_{2.5}: $r > 0.95$
 - Temporal correlations of satellite and ground PM_{2.5}: $r = 0.81\text{-}0.87$

□ Emergency department visit data

- Individual-level ED visits obtained from 40 acute care hospitals in the study area
- ED daily counts of respiratory outcomes aggregated by grid cell

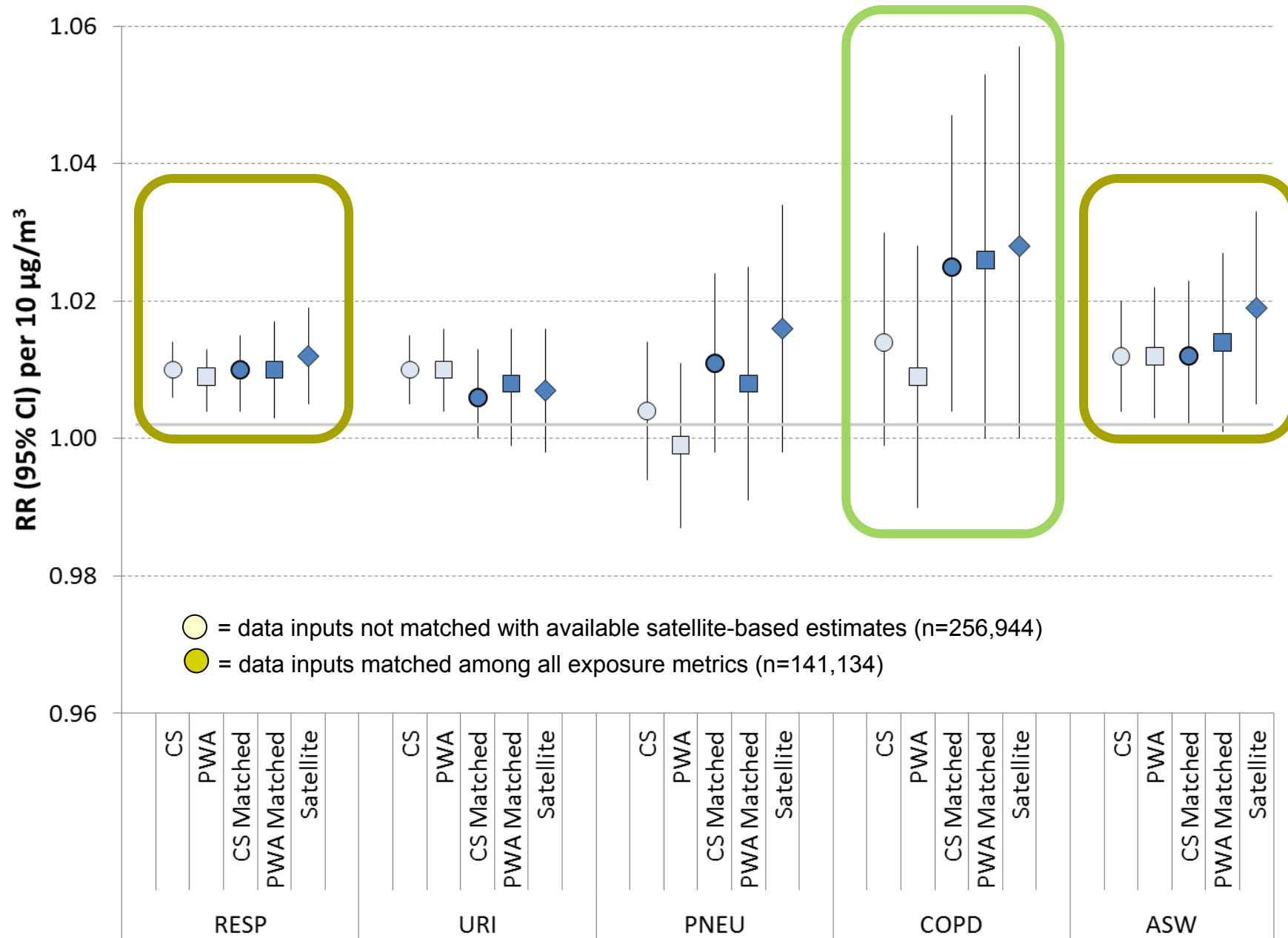
Outcome		ED visits per day		ED visits per day per grid cell	
		Mean	Std Dev	Mean	Std Dev
Respiratory disease	RESP	386.1	131.7	3.8	5.8
Upper respiratory infection	URI	226.3	84.4	2.2	3.5
Pneumonia	PNEU	49.6	19.9	0.5	0.9
COPD	COPD	18.4	6.7	0.2	0.5
Asthma/Wheeze	ASW	78.3	27.1	0.8	1.7

- Poisson generalized linear models to estimate the association between ED visits and same-day $PM_{2.5}$

$$E(Y_{gt}) = \beta_0 + \beta_1 PM_{gt} + \text{covariate terms}$$

- Y_{gt} represents the ED counts for grid g on day t
- Hybrid case-crossover smoothing approach, where covariate are:
 - Indicator variable for grid cell → controls for spatial confounding, such as SES
 - Indicator variables for month, year, day-of week/holidays & interactions
 - Spline terms with seasonal knots but with the same parameters each year
 - Cubic terms for max temp and its linear interaction with month
 - Cubic terms for min temp and cubic terms for mean dew point
 - Hospital indicators if they don't coincide with the month dummy variables

Epidemiologic Results



- ❑ Impact of spatial refinement comparing CS, PWA, and satellite matched results
 - ❑ RRs among the 3 exposure metrics consistent (confidence intervals all overlap)
 - ❑ Trends for slightly stronger associations using satellite estimates; suggestive of reduced exposure measurement error
- ❑ Population density highest towards center of study domain, therefore general agreement of CS and satellite-based RRs anticipated
- ❑ 12 km grid data may be too coarse to see much more refinement in exposure precision